CpE 645 Image Processing and Computer Vision

Prof. Hong Man

Department of Electrical and Computer Engineering
Stevens Institute of Technology

STEVENS Institute of Technology

Visual Information Environment Laboratory

Representations & Features

- The segmentation techniques yield raw image data in the form of pixels along a boundary or pixels contained in a region.
- Representation schemes are used to compact the data into representations that are considerably more useful in the computation of descriptors (or features).
- Representing a region involves two choices:
  - in terms of its boundary (shape)
  - in terms of its internal characteristics (texture, color)
- The features selected as descriptors should be as insensitive as possible to variations such as changes in size, translation, and rotation.

Chain-Code Representation

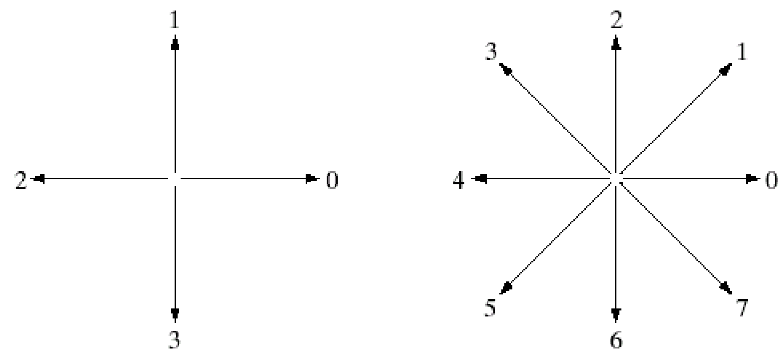
- Boundary chain codes are representations that can effectively trace out the boundary of a binary object.
- They exploit the property that the boundary path is connected.
- The procedure :
  - Select a point on the boundary of the object as an initial starting point, and a direction for the path (such as clockwise).
  - Examine the four or the eight surrounding neighbors to determine the next point along the clockwise boundary path.
  - Express the direction number in binary code.

Chain-Code Representation

- The 4-connected and the 8-connected pixels

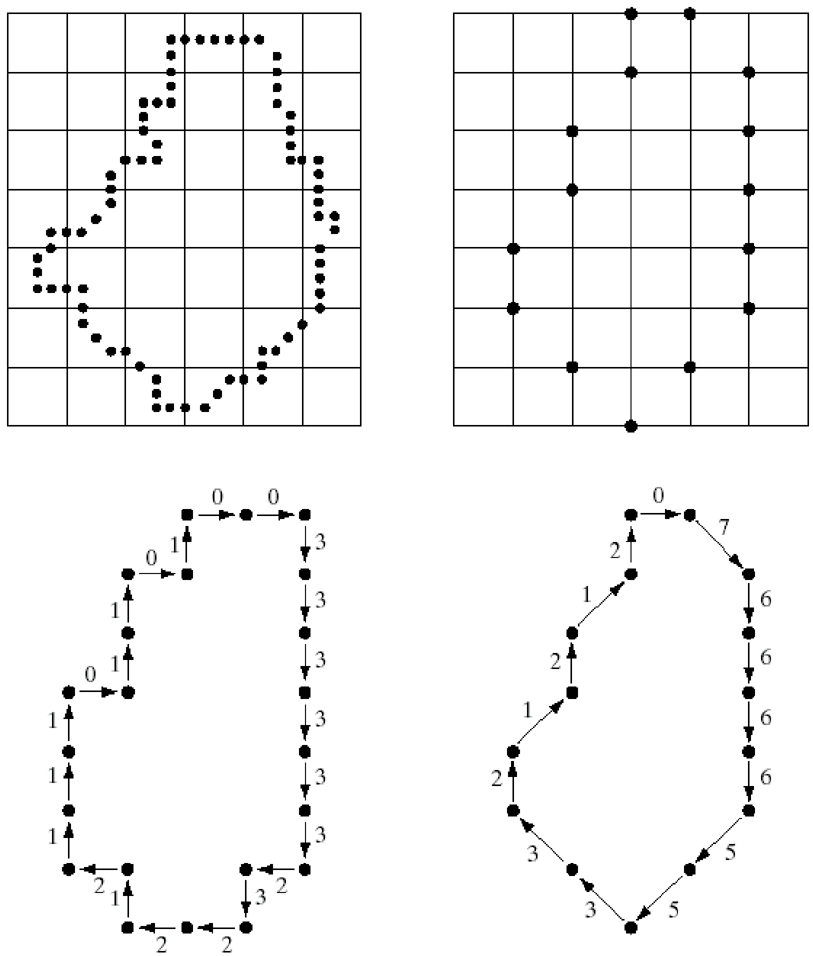
a b

FIGURE 11.1 Direction numbers for (a) 4-directional chain code, and (b) 8-directional chain code.



- Thus, the entire chain-code representation consists of the initial starting coordinates (n_1, n_2) and a sequence of binary codes denoting the directions of each step as the boundary path is traversed, e.g 2-bit codes for 4-connected and 3-bit codes for 8-connected.

Chain-Code Representation



a b
c d

FIGURE 11.2 (a) Digital boundary with resampling grid superimposed. (b) Result of resampling. (c) 4-directional chain code. (d) 8-directional chain code.

Chain-Code Representation

- If start from top-middle point,
  - the chain code for (c): 00333…01, or in binary
    00 00 11 11 11 … 00 01
  - the chain code for (d): 07666…12, or in binary
    000 111 110 110 110 … 001 010
- The chain code can be normalize w.r.t. the stating point by using the *first difference* of the chain code. The difference is obtained simply by counting (counter-clockwise) the number of direction changes that separate two adjacent elements of the code.
  - Example: the first difference of the 4-connected chain code 10103322 is 3133030

Shape Numbers

- *Shape numbers* can be used for shape description. It is defined as the first difference with smallest magnitude.
- The digit sequence of the first difference is a circular sequence depending on the starting point.
- Shape number is formed by rotating the first difference sequence and finding a starting point that yields a sequence of digits, if collectively considered as <u>one integer</u> number, that has the smallest value.
- The order **n** of a shape number is defined as the number of digits in its representation.
- The order **n** is always even for a closed boundary.

Shape Numbers

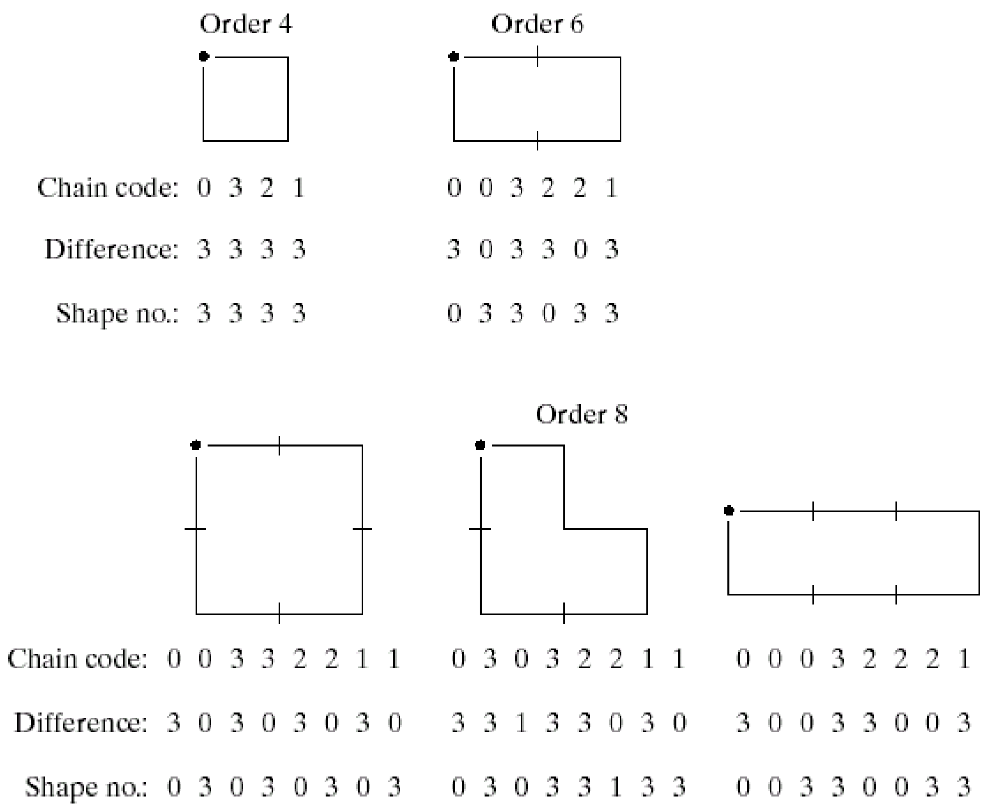
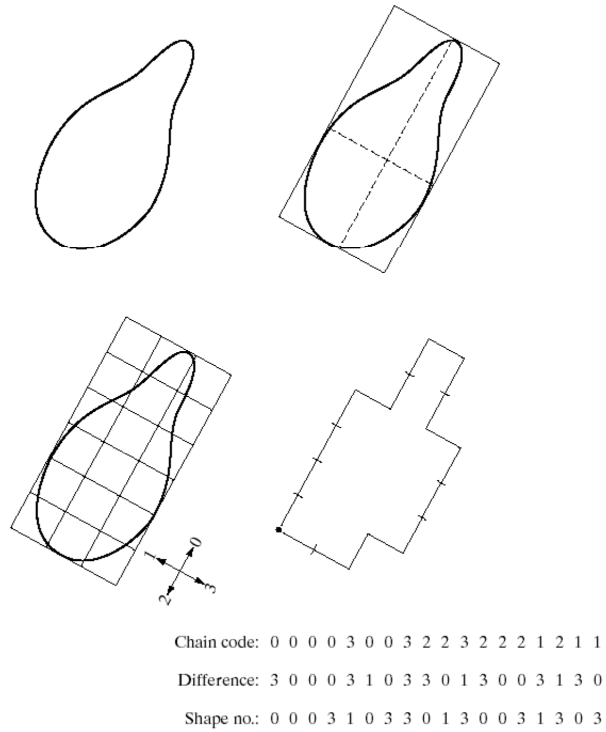


FIGURE 11.11 All shapes of order 4, 6, and 8. The directions are from Fig. 11.1(a), and the dot indicates the starting point.

Shape Numbers

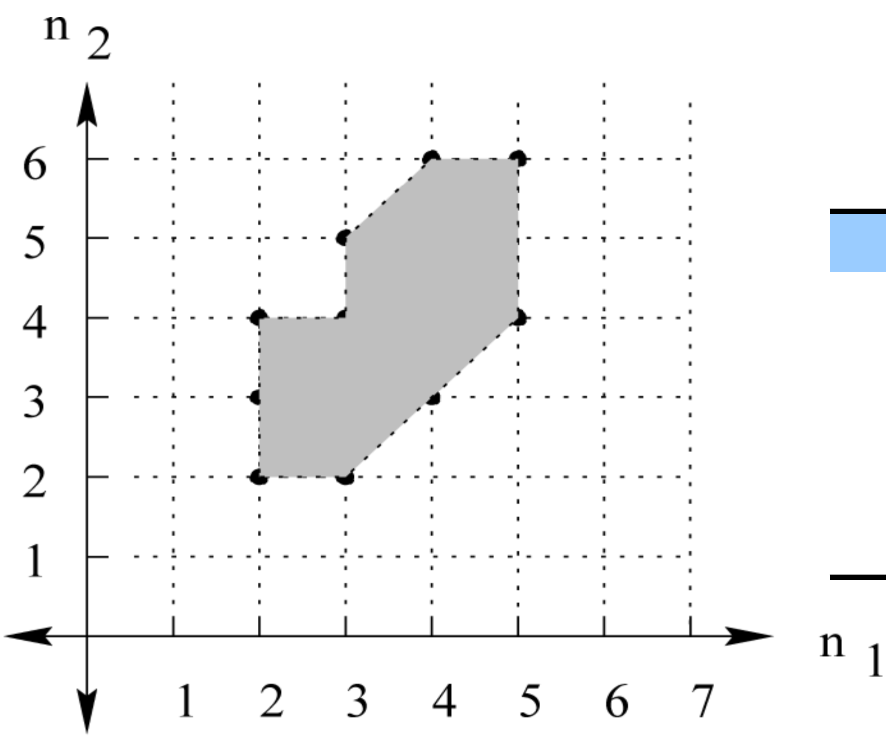
- To obtain an order **n** shape number, we need to find a basic rectangle with block number of **n.** Each such block will produce one segment of the boundary. (**n=18** in the example)



Run-Length Representation

- The object shape can be viewed as a binary image, therefore *run-length coding* can be applied to represent this object.
- This procedure scans over all the horizontal (or vertical) lines. At each line, find the starting point and recording consecutive <u>runs of zeros or ones</u> on that line.
- The form of the code is a sequence of three letter triplets **(n_1, n_2, run-length),** where n_1, n_2 are the starting coordinates.
- Runs can be ones if it is a bright object on a dark background, or can be zeros if a dark object on a bright background.

Run-Length Representation

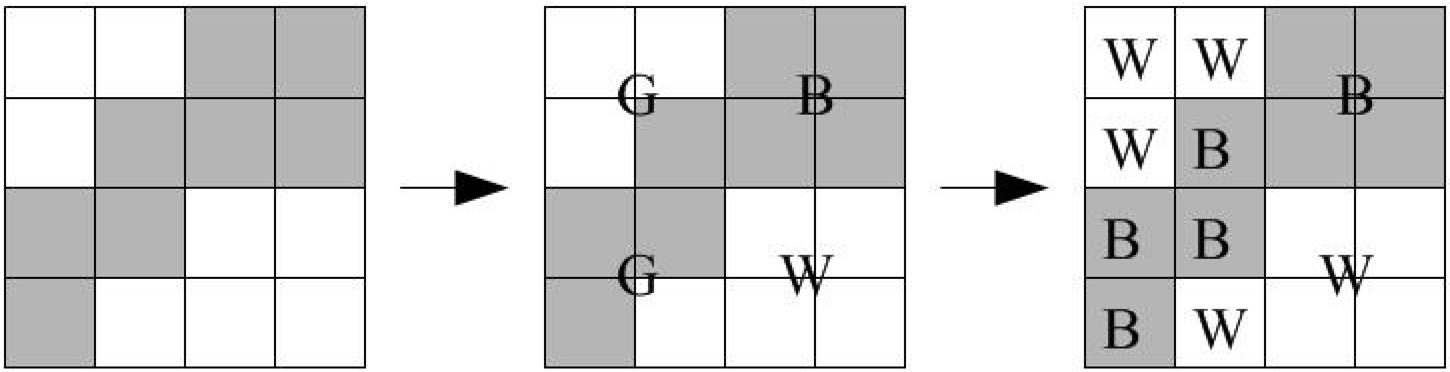


n1	*n2*	run-length
4	6	1
3	5	2
2	4	3
2	3	2
2	2	1

Quad-Tree Representation

- Quad-trees can be used to represent a binary region or object by successive divisions of the image into four quadrants. The method employs a set of three symbols in its representation and uses a simple splitting rule.
- Three symbols:
  - B: all black pixel quadrant
  - W: all white pixel quadrant
  - G: mixed black and white pixel quadrant.
- The splitting rule consists of successive splitting of all gray blocks until the entire image is composed solely of black and white blocks.

Quad-Tree Representation



The resulting quad-tree code is:

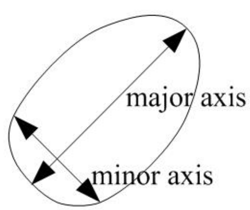
G,B,G,W;W,W,W,B;B,B,B,W

Shape Descriptors

- Properties of desirable shape descriptors:
  - Invariance: to linear transformations (rotation, scaling, translation)
  - Completeness: (the descriptor should identify the shape uniquely)
- Important simple shape descriptors:
  - Geometric descriptors: length, eccentricity, roundness, curvature
  - Fourier descriptors
  - Statistical moments

Shape Descriptors

- Length (perimeter): compute the number of ones on the boundary
- Eccentricity: ratio of the major and minor axes
- Roundness or compactness: **(perimeter)^2/(4π×area)**
- Curvature: the rate of change of slopes along the boundary



eccentricity

curvature

Fourier Descriptors

- The coordinate pairs on the boundary, **(x[0], y[0]), (x[1], y[1]), (x[2], y[2]) …(x[N-1], y[N-1]),** can be considered as a complex waveform in the term of

$$u[n]=x[n]+jy[n]$$

- The DFT of signal $u[n]$ is

$$a[k]=\sum_{n=0}^{N-1}u[n]e^{\frac{-j2\pi kn}{N}}$$

- The complex coefficients $a[k]$ are called Fourier descriptors (FDs) of the boundary

Fourier Descriptors

- Geometrical transformations of a boundary can be related to simple operations on the FDs

Transformation	Boundary	Fourier Descriptor
Identity	$s(k)$	$a(u)$
Rotation	$s_r(k) = s(k)e^{j\theta}$	$a_r(u) = a(u)e^{j\theta}$
Translation	$s_t(k) = s(k) + \Delta_{xy}$	$a_t(u) = a(u) + \Delta_{xy}\delta(u)$
Scaling	$s_s(k) = \alpha s(k)$	$a_s(u) = \alpha a(u)$
Starting point	$s_p(k) = s(k - k_0)$	$a_p(u) = a(u)e^{-j2\pi k_0 u/K}$

TABLE 11.1 Some basic properties of Fourier descriptors.

Fourier Descriptors

- The FD magnitudes have some invariant properties
  - $|a[k]|$ for **k**=1, …, N-1 is invariant to starting point, rotation, and reflection.
  - $a[k]/|a[k]|$ is invariant to scaling
  - note: magnitude along can not reconstruct the original boundary.

Fourier Descriptors

- FD can be truncated (stop at k<N-1) to produce an approximation of the boundary.

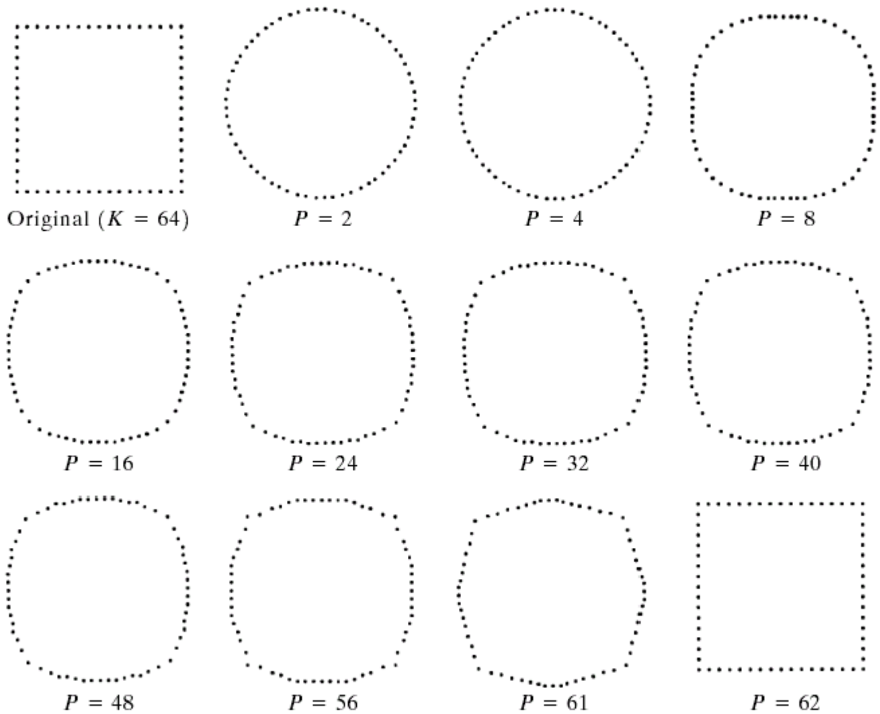


FIGURE 11.14 Examples of reconstruction from Fourier descriptors. P is the number of Fourier coefficients used in the reconstruction of the boundary.

Moment Representations

- Moments can be used as texture descriptors.
- Let $f(x, y)$ be a positive real bounded continuous image with finite support in the region R. The $(p+q)^{th}$-order moment is defined as,

$$m_{pq} = \iint_R f(x, y) x^p y^q \, dx dy$$

  where p and q are non-negative integers.
- The center of gravity (or centroid) of the object is given as

$$\bar{x} = \frac{m_{10}}{m_{00}} \qquad \bar{y} = \frac{m_{01}}{m_{00}}$$

- And the central moments are defined as

$$\mu_{pq} = \iint_R f(x, y)(x - \bar{x})^p (y - \bar{y})^q \, dx dy$$

Moment Representations

- For digital images, the moment is

$$m_{pq} = \sum_i \sum_j i^p j^q f[i, j]$$

- For binary digital images, the moment reduces to

$$m_{pq} = \sum_i i^p \sum_j j^q$$

  and m_{00} is simply the total number of points in the region. The centroid is given by m_{01} m_{10}

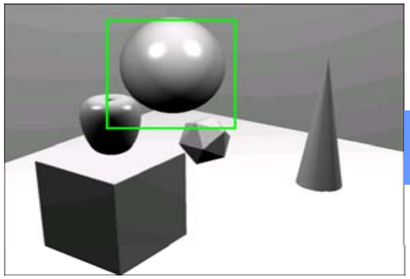
- High order moments and central moments are sensitive to scaling and rotation.

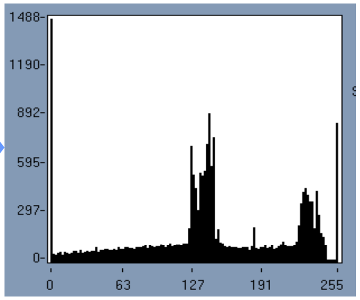
Histogram Features

- A set of features derived from the local histogram can be used for texture descriptor.
- The histogram and all resulting features are based on a localized spatial window of size $\boldsymbol{M \times M}$, where $\boldsymbol{M \times M}$ is much smaller than the image size.
- Consider histogram $\boldsymbol{P_x[l]}$ as the probability density function of $\boldsymbol{x}$. A set of common parametric descriptors can be defined:
  - mean $$\overline{m} = \sum_{l=0}^{L-1} l P_x[l]$$
  - Standard derivation $$\sigma_x^{\ 2} = \left(\sum_{l=0}^{L-1} (l - \overline{m})^2 P_x[l] \right)^2$$

Histogram Features

- Skewness $$S = \frac{1}{\sigma_x^3}\sum_{l=0}^{L-1}\left(l-\overline{m}\right)^3 P_x[l]$$
- Kurtosis $$K = \frac{1}{\sigma_x^4}\sum_{l=0}^{L-1}\left(l-\overline{m}\right)^4 P_x[l]-3$$
- Energy $$E = \sum_{l=0}^{L-1}\left(P_x[l]\right)^2$$
- Entropy $$H = -\sum_{l=0}^{L-1} P_x[l]\log_2\left(P_x[l]\right)$$





Classification

- Classification is to determine the type of an object.
- The object under evaluation is represented by a set of features (i.e. the feature vector):

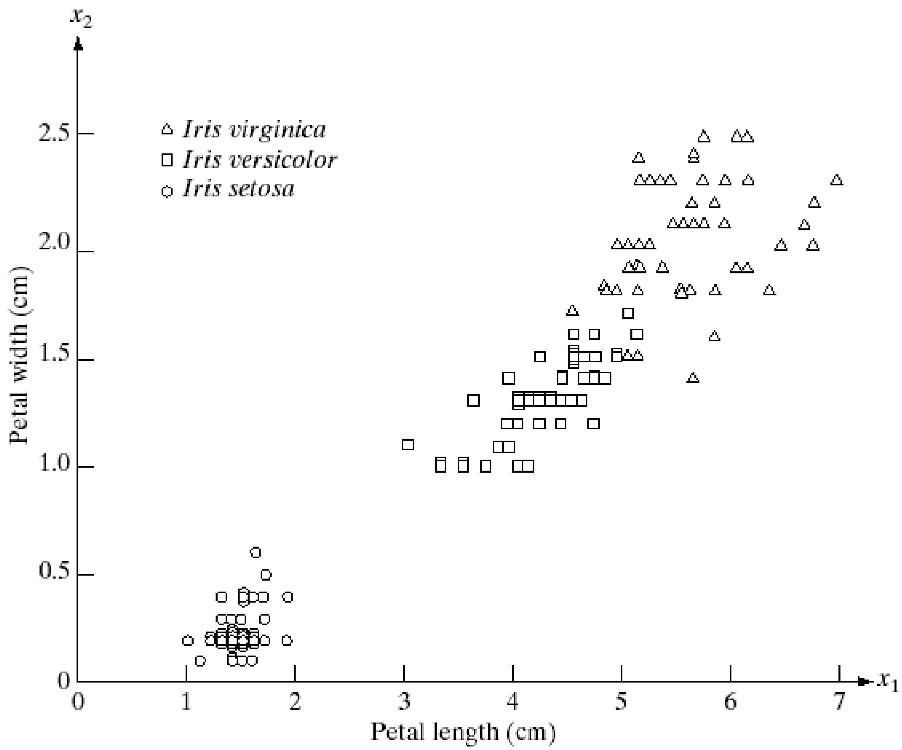
$$\boldsymbol{x} = [x_1, x_2, \ldots, x_N]^T$$

- The classification procedure
  - For the input scene, a feature vector is formed.
  - The classifier determines to which of the M possible classes the feature vector belongs.

Classification

- Example: describe three types of iris flowers by measuring the widths and lengths of their petals

FIGURE 12.1
Three types of iris flowers described by two measurements.



Classification

- Classification is based on the use of decision functions.
- The basic problem is to find M decision functions $d_1(x), \ldots, d_M(x)$ with the property that if a pattern x belongs to class c_i, then

$$d_i(x) > d_j(x) \qquad \forall j = 1,\ldots,M, \quad j \neq i$$

- The decision boundary separating class c_i from c_j is given by the values of x for which $d_i(x)-d_j(x)=0$
- The objective is to develop various approaches to find effective decision functions

Minimum Distance Classifier

- Assume each pattern class is represented by a prototype (mean) vector $\boldsymbol{m_j}$
- One way is to assign $\boldsymbol{x}$ to the class of its closest prototype is to compute the distance measure

$$D_j(x) = \|x - m_j\| \quad j = 1,\dots,M \qquad \text{where} \quad \|a\| = (a^T a)^{1/2}$$

  and select the minimum distance.
- This is equivalent to evaluate the functions

$$d_j(x) = x^T m_j - \frac{1}{2} m_j^T m_j \qquad j = 1,\dots,M$$

  and assign $\boldsymbol{x}$ to class $\boldsymbol{c_i}$ if $\boldsymbol{d_i(x)}$ is maximum.
- So the decision boundary between classes $\boldsymbol{c_i}$ and $\boldsymbol{c_j}$ for a minimum distance classifier is:

$$d_{ij}(x) = d_i(x) - d_j(x) = x^T(m_i - m_j) - \frac{1}{2}(m_i - m_j)^T(m_i - m_j) = 0$$

Minimum Distance Classifier

- Note that the surface of the decision boundary is perpendicular bisector of the line segment joining $\boldsymbol{m}_i$ and $\boldsymbol{m}_j$

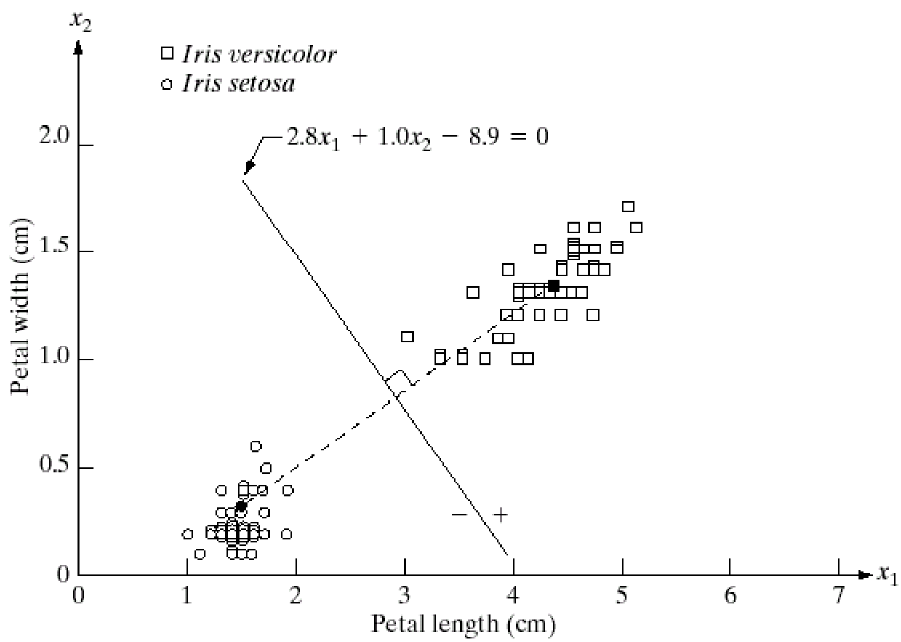


FIGURE 12.6 Decision boundary of minimum distance classifier for the classes of *Iris versicolor* and *Iris setosa*. The dark dot and square are the means.

Minimum Distance Classifier

- The minimum distance classifier works well when the distance between means is large compared to the spread or randomness of each class with respect to its mean.
- However, the occurrence of large mean separations and relatively small class spread occur seldomly.

Matching By Correlation

- Match by correlation intends to find matches of a subimage w(x,y) in an image f(x,y).
- Assume the size of w(x,y) is J×K, and size of f(x,y) is M×N, and J≤M, K≤N. The correlation is calculated as

$$c(x,y) = \sum_s \sum_t f(s,t)w(x+s, y+t)$$

for x = 0, 1, …, M-1,
and y = 0, 1, …, N-1.

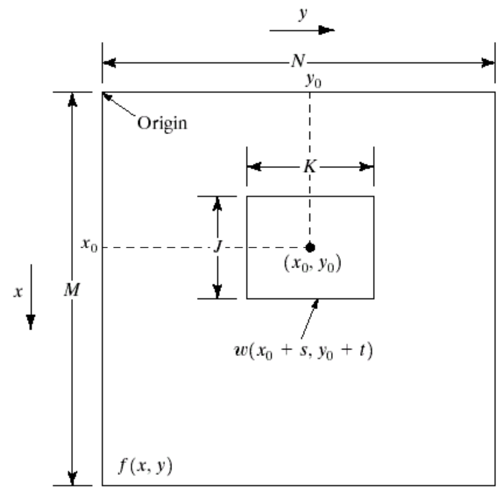


FIGURE 12.8 Arrangement for obtaining the correlation of f and w at point (x_0, y_0).

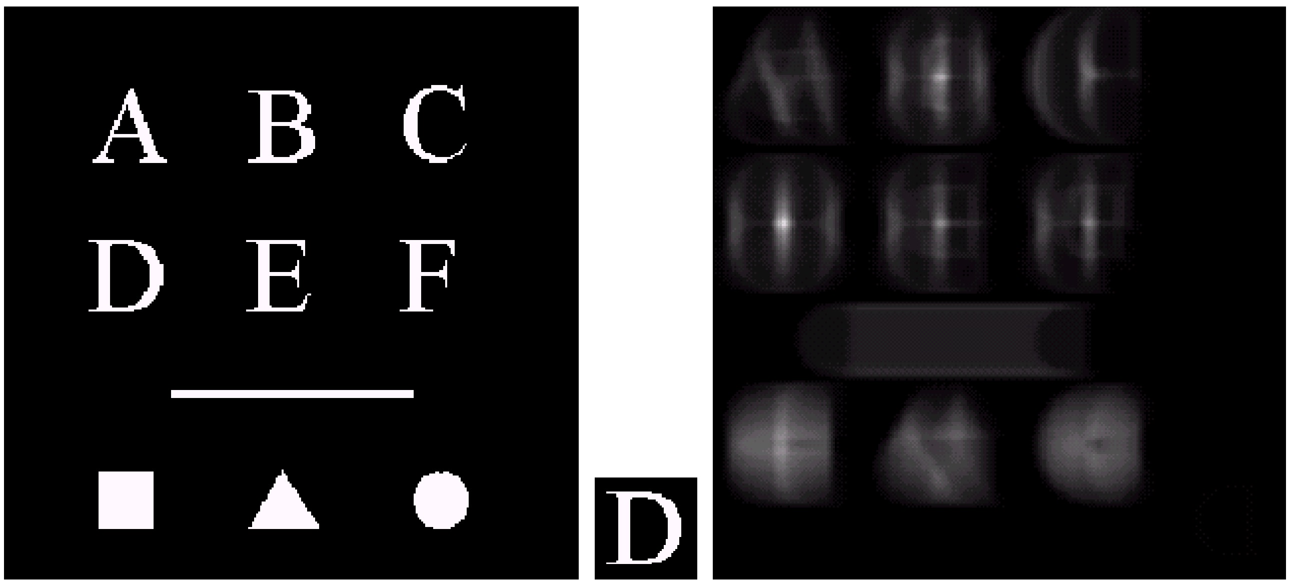
Matching By Correlation

- To normalized correlation w.r.t. image amplitude, correlation coefficient is defined as

$$\gamma(x,y)=\frac{\sum_s\sum_t\left[f(s,t)-\bar{f}(s,t)\right]\left[w(x+s,y+t)-\bar{w}\right]}{\left\{\sum_s\sum_t\left[f(s,t)-\bar{f}(s,t)\right]^2\sum_s\sum_t\left[w(x+s,y+t)-\bar{w}\right]^2\right\}^{1/2}}$$

where $\bar{w}$ is the average of the subimage, and $\bar{f}$ is the average of $f(x,y)$ inside the w window.

Matching By Correlation



a b c

FIGURE 12.9
(a) Image.
(b) Subimage.
(c) Correlation coefficient of (a) and (b). Note that the highest (brighter) point in (c) occurs when subimage (b) is coincident with the letter "D" in (a).

Optimum Statistical Classifier

- It is possible to derive a classification method that is optimal in the sense that on average it has the lowest probability of causing classification errors.
- Let $p(c_i/x)$ denote the probability that a particular pattern x comes from class c_i out of M classes.
- If the classifier decides that x comes from class c_j, when it is actually comes from class c_i, it incurs a loss denoted by L_{ij}. The average loss (also called the *conditional average risk*) in assigning x to c_j is then

$$r_j(x) = \sum_{k=1}^{M} L_{kj} p(c_k/x)$$

Optimum Statistical Classifier

- According to the Bayes' rule,

$$r_j(x) = \frac{1}{p(x)} \sum_{k=1}^{M} L_{kj} p(x/c_k) p(c_k)$$

- The classifier that minimizes the total average loss is called the Bayes classifier; that is it assigns an unknown pattern $\boldsymbol{x}$ to class c_i if $r_i(\mathbf{x}) < r_j(\mathbf{x})$ or

$$\sum_{k=1}^{M} L_{ki} p(x/c_k) p(c_k) < \sum_{q=1}^{M} L_{qj} p(x/c_q) p(c_q)$$

- The loss function can be defined as $L_{ij} = 1 - \delta_{ij}$ which indicates a loss of **1** for incorrect decision and a loss of **0** for correct decisions.

Optimum Statistical Classifier

- With this L_{ij}, the decision is made based on

$$\sum_{k=1}^{M}(1-\delta_{ik})p(x/c_k)p(c_k) < \sum_{q=1}^{M}(1-\delta_{qj})p(x/c_q)p(c_q)$$

then it becomes

$$p(x) - p(x/c_i)P(c_i) < p(x) - p(x/c_j)P(c_j)$$

or

$$p(x/c_i)P(c_i) > p(x/c_j)P(c_j)$$

for $j = 0, 1, \ldots, M$, and $j \neq i$

so the Bayes classifier for a 0-1 loss function is the computation of the decision function

$$d_j(x) = p(x/c_j)P(c_j)$$

Optimum Statistical Classifier

- Bayes classifier require that the probability density functions of the patterns in each class $p(x/c_i)$, as well as the probability of occurrence in each class $P(c_i)$ must be known.
- Estimation of $p(x/c_i)$ can be a problem. This requires methods from multivariate probability theory. Difficult to apply in practice.
- The use of Bayes classifier is generally based on the assumption of an analytic expression of the density functions and an estimation of the parameters from sample patterns for each class.

Clustering

- After a measure of the similarity of the testing patterns, a procedure has to be specified to partition the given data into cluster domains
- Simple cluster seeking (SCS) algorithm:
  - Specify a distance measure and a distance threshold $\boldsymbol{T}$,
  - Arbitrarily select a pattern representation vector $\boldsymbol{x}_1$, and assign it as the center $\boldsymbol{z}_1$ of the first cluster $\boldsymbol{S}_1$
  - Take the next vector $\boldsymbol{x}_2$, and measure the distance between $\boldsymbol{z}_1$ and $\boldsymbol{x}_2$, i.e. $\boldsymbol{D}_1(\boldsymbol{x}_2)$
  - If $\boldsymbol{D}_1(\boldsymbol{x}_2) \leq \boldsymbol{T}$, assign $\boldsymbol{x}_2$ to the first cluster $\boldsymbol{S}_1$,
  - Otherwise, assign $\boldsymbol{x}_2$ as the center $\boldsymbol{z}_2$ of the second cluster $\boldsymbol{S}_2$

Clustering

- In general, take any not yet clustered vector $\boldsymbol{x}_k$, and measure the distance between $\boldsymbol{z}_i$ and $\boldsymbol{x}_k$, i.e. $\boldsymbol{D}_i(\boldsymbol{x}_k)$, for all $\boldsymbol{i}=\mathbf{1}\ldots\boldsymbol{N}$, where N is current number of all clusters
- If $\boldsymbol{D}_i(\boldsymbol{x}_k) \leq \boldsymbol{T}$, assign $\boldsymbol{x}_k$ to cluster $\boldsymbol{S}_i$
- Otherwise, assign $\boldsymbol{x}_k$ as the center $\boldsymbol{z}_{N+1}$ of a new cluster $\boldsymbol{S}_{N+1}$, and let $\boldsymbol{N}=\boldsymbol{N}+\mathbf{1}$.
- This procedure continues until all vectors have been assigned to one cluster.

Clustering

- Notes on SCS:
  - The results of this algorithm depend on the first cluster center chosen, the order in which the patterns are considered, the value of $\boldsymbol{T}$, and the data itself.
  - A vector is not assigned to the cluster with minimum distance, instead, it is assigned to a cluster that first satisfies $\boldsymbol{D_i(x_k) \leq T}$ according to the testing order.
  - In practice, this procedure normally requires extensive experimentation with various values of the threshold and different starting points.

Clustering

- K-means clustering
  - Select k initial cluster centers $z_1(1), z_2(1), \ldots z_K(1)$.
  - At the k-th iterations, distribute all the vectors$\{x\}$ among the K clusters according to

$$x \in S_i(k) \ \text{if} \ \|x\text{-}z_i(k)\| < \|x\text{-}z_j(k)\|$$

$$\text{for all } i = 1, \ldots K \text{ and } i \neq j$$

  - Computer the new cluster centers according to

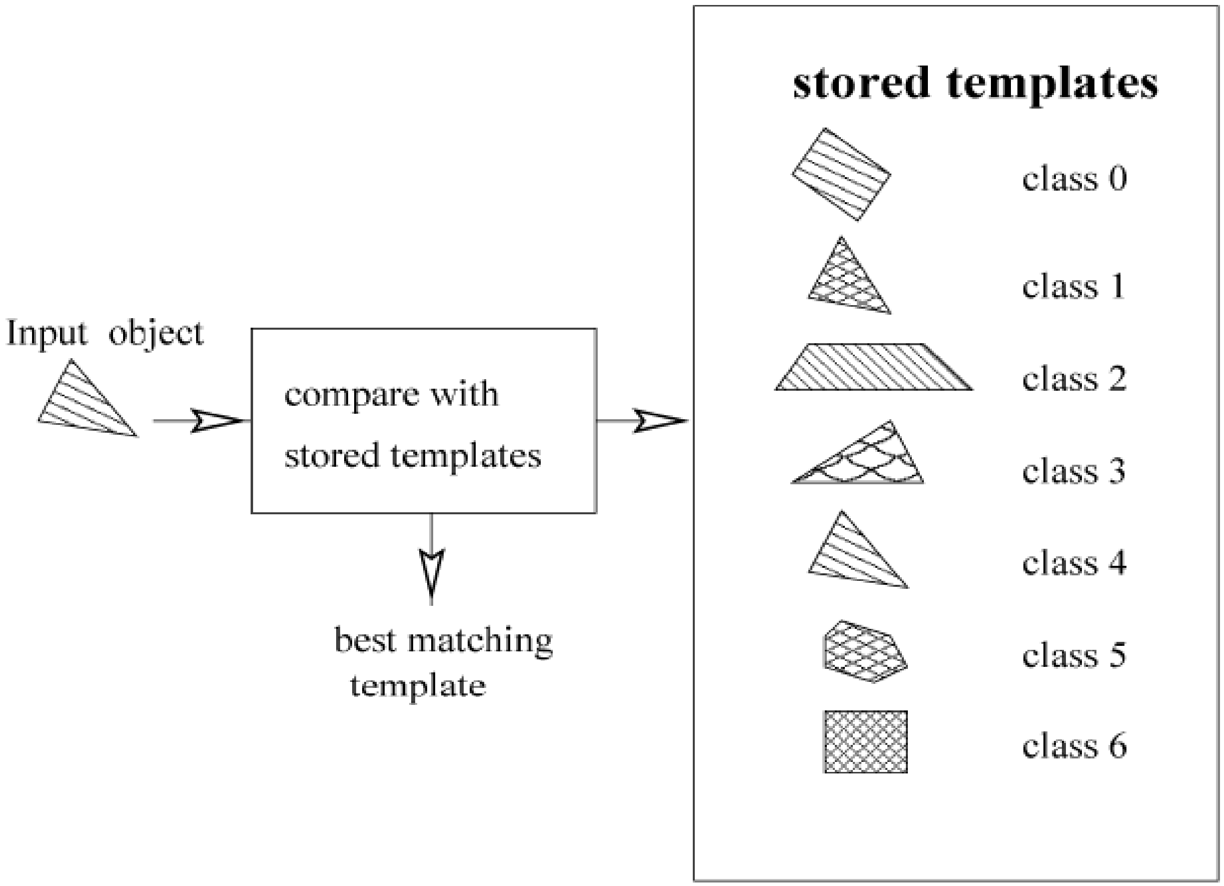
$$z_i(k+1) = \frac{1}{N_i} \sum_{x \in S_i(k)} x \ \text{ for } i = 1, \ldots K$$

  - The algorithm converges if $z_i(k+1) = z_i(k)$
  - Otherwise go to iteration k+1

Template Matching

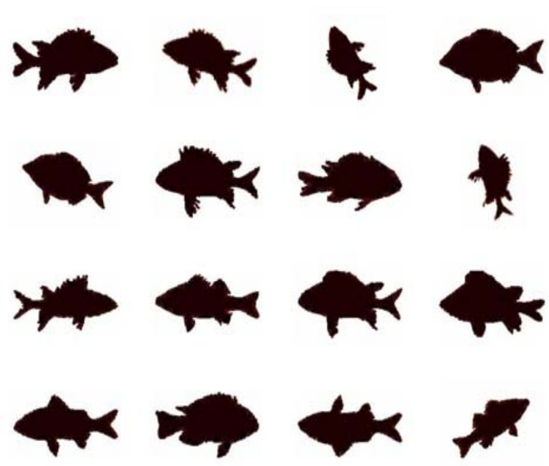
- Template matching is the most intuitive method of classification. The classifier stores a set of templates in its memory. Each template has a class associated with it.
- Each input object (representation vector) is compared to each of every template stored in memory. The class associated with the template producing the smallest distance is then recorded as the class of the object.
- The distance measure normally used in template matching can be mean square error or mean absolute error. These can also be calculated based on vectors, i.e. norms.

Template Matching



Template Matching

- For reliable performance over a wide class of inputs, it is generally necessary to have several templates associated with a given class. These templates may differ in terms of their orientation, position, and size.



Object Detection

- A broad class of image analysis problems involve detecting the presence of an object or target in an image.
- Example:
  - Detecting military targets in forward-looking infrared (FLIR) images. These sensors produce images where warm or hot regions show up as being bright, while regions that are cold appear dark.
  - Synthetic aperture radar (SAR). These images will often display intense bright spots at object corners, where the radar returns are direct and strong. A popular detection problem for SAR images is to find military targets. Such targets (as well as other man-made objects) will generally contain very bright spots.

Object Detection

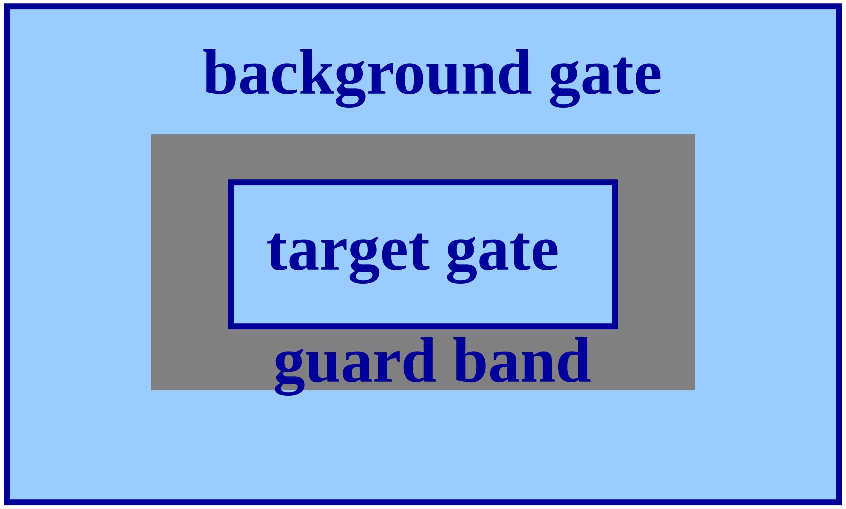
- Detection of abnormalities in x-ray images. If a tumor is present, one generally sees small regions of discoloration.
- Manufacturing: Automated inspection systems can improve quality control and save money.

Contrast Box Method

- The contrast box method is based on using the relative amplitude differences between the target and the surrounding background as discriminate features.
- The amplitude difference is measured within a box-shaped region of support, where the inner box (a.k.a. the target gate) is chosen to be about the same size as the target. The dimensions of the outer box are chosen to be larger than the target, making its region of support appropriate for estimating the background.
- Since the dimensions of the targets and the range (distance) to the targets are generally known in these situations, this information can be used to set the contrast box dimensions.

Contrast Box Method

- The contrast box works as a filter. That is, its position over the image is moved in increments of one pixel at a time. However, instead of multiplying and summing at each position (as in convolution), the mean and variance within the regions are computed.



Contrast Box Method

- Let μ_T and σ_T^2 denote the mean and variance of the target box, and μ_B and σ_B denote the mean and variance of the background gate.
- The typical contrast-box metric that is calculated at each pixel position is

$$C_{CB} = \frac{(\mu_T - \mu_B)^2 + \sigma_T^{\ 2}}{\sigma_B}$$

- The output of this operation is a gray-scale image containing the values of C_{CB}, computed at each pixel location in the image. High amplitudes of C_{CB} indicate high regional statistical deviation, which usually imply regions of interests.

Contrast Box Method

- A popular variation of contrast box method is the constant false alarm rate technique (CFAR).
- In CFAR,
  - The small box in the center (the *test cell*) is typically a single pixel or a small region of size equal to or smaller than the target.
  - A guard-band is imposed around the test cell to assure that the possible target region does not intersect the regions for background estimation.
  - The lagging and leading background cells play the role of the background gate

Contrast Box Method

- Detection is performed by comparing the image amplitude in the test cell with a threshold, derived from the background cells.
- A target is said to be present if the test cell pixel exceeds the threshold.